UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/581,290	06/01/2006	Simon PA Ringland	36-1990	3165

23117   7590   05/04/2010
NIXON & VANDERHYE, PC
901 NORTH GLEBE ROAD, 11TH FLOOR
ARLINGTON, VA 22203

EXAMINER
BORSETTI, GREG

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
05/04/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on <u>30 March 2010</u>.

2a) ☒ This action is **FINAL**.   2b) ☐ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) <u>19-27 and 29-36</u> is/are pending in the application.

     4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) <u>19-27 and 29-36</u> is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☒ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

     a) ☒ All   b) ☐ Some * c) ☐ None of:

        1. ☒ Certified copies of the priority documents have been received.

        2. ☐ Certified copies of the priority documents have been received in Application No. _____.

        3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

     * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application.

6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1.   Claims 19-27, and 29-36 are pending.

2.   Claims 19, 27, and 29 have been amended.

3.   Claim 28 has been canceled, therefore the objection to claim 28 has been

withdrawn.

4.   The 35 USC 112 rejections to claims 27-28 have been withdrawn in view of the

amendments (3/30/2010).

Response to Arguments

5.   Applicant argues "Applicants' claimed invention differs from Kinnunen, inter alia,

in that the voice recognition system is incorporated in a server in the network (which is

now clearly stated in the independent claims). This has the advantage that any push-to-

talk device even a simple device without voice recognition features, can be used in the

system. Further, the prior art problem of too much power consumption in mobile

telephones having a speech recognition engine is avoided with this solution." (Remarks,

Page 11, ¶ 2) In response to applicant's argument that the references fail to show

certain features of applicant's invention, it is noted that the features upon which

applicant relies (i.e., ...*in a speech recognition server in said communications*

network...) are not recited in the previously rejected claim(s). Although the claims are

interpreted in light of the specification, limitations from the specification are not read into

the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

6.   Applicant further argues "Finally, the purpose of Kinnunen is not to provide a push-to-talk system that does not require a separate dialing phase, but rather to find a solution that reduces the high power consumption of mobile devices having a speech recognition engine." (Remarks, Page 11, ¶ 2) Applicant's arguments fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references. Furthermore, applicant, in attempting to recognize the inventive concept behind Kinnunen, fails to recognize what Kinnunen actually performs in operation. The argument is not persuasive.

7.   Applicant further argues "With respect, whether or not a person of common sense addresses a target individual before speaking to them is irrelevant. Besides the fact that many people speak to other people without having prefaced such speaking with some sort of "address", this has nothing to do with either teaching or suggesting that a speech recognition process is to be performed on only a portion of a received audio stream - when the intended recipient is indicated at the beginning of the audio stream. The asserted rationale for finding "obviousness" is simply a non sequitur." (Remarks, Page 12, ¶ 2) The Examiner disagrees. First, the speech recognition (VRE) is only performed on a portion of the audio stream because the received audio is continuous (noise/speech), but the speech recognition is only performed on speech (a portion). Second, the intended recipient is taught by the password on Page 17, lines 11-24. Applicant chooses to argue the differences between "password" and "keyword" in the subsequent paragraph, but it will be addressed here. Page 17, lines 19-24 teach

"This may also be implemented with the VAD and VRE modules of a basic model presented above, wherein the VAD module detects starting and ending points of a sentence and the VRE module recognizes a keyword and the transmission is controlled not by pushing/releasing a tangent but according to starting and ending points of a sentence detected by the VAD module." This shows that the "password" enacted through the VAD and VRE creates the hand-free mode discussed in the same paragraph where according to Page 17, lines 26-34, the message and recipient may all be within one "dialing phase". Furthermore, the Examiner still contends that it would have been obvious to someone of ordinary skill in the art that the "password" could have been the name of the individual the user wishes to connect to because of the reason set forth in the office action, and further because of Page 17, lines 26-34 where the system identifies a recipient in the message stream and the system send the message to that person. Claims 7-8 further show this in that the VRE function searches for a keyword, recognizes a keyword, and uses that keyword to determine the recipient (all performed in a single step and vocally). In this scenario, there are three options. Either the recipient is named before, after, or in the middle of the message which at least teaches the "the intended recipient is indicated at the beginning of the audio stream". This at least makes Kinnunen obvious over claim 23. Furthermore, Kinnunen, Page 28, lines 15-29 teaches that there is a confirmation of the recipient before the message. This confirmation is not a separate dialing phase and thus further teaches over claim 23. The argument is not persuasive.

8.   Applicant further argues "However, the Vysotsky teaching at 5:45-50 requires a

voice verification circuit 255 when it is important to verify the entity of a caller before responding to a particular command. This is part of a voice recognition circuit 250. It would appear that only something downstream from the voice recognition circuit 250 might be "receiving an indication of the identity of a user who generated the audio stream". The generation of a voice identification of a user as taught at 5:45-50 is hardly a means for receiving an indication of the identity of a user, etc.

The passage at 8:31-35 of Vysotsky deals with speaker-dependent speech recognition processes based on hidden Markov models with the use of "grammars". It appears to deal with step 403 depicted in Fig. 4 - which it will be noted is upstream of arbitration 406 and call completion 424 (which logically appears to be associated with the "call completion and feature activation" block 256 back in Fig. 2B that has some relationship to the earlier cited passage at 5:45-50)" (Remarks, Page 13, ¶ 3 - ¶ 4) Applicant's arguments fail to comply with 37 CFR 1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references.

9.   Applicant further argues "In any event, even though the Examiner has attempted to splice two disparate portions of the Vysotsky teaching, even if they are illogically "spliced", they still do not teach selecting a user-dependent speech grammar for use by a speech recognition process in dependence upon the identity of the user- which was previously indicated by receiving an indication of the identity of a user who generated the audio stream, etc. Where, for example, is there any possible suggestion in Vysotsky

of using a different grammar in a speech recognition process depending on the identity

of the person who generated the audio stream? The Examiner's hypothesis that "there

would be a grammar selection means for selecting a user-dependent speech grammar

dependent for the specific user if voice verification is performed on the individual" is

merely speculation by the Examiner. Furthermore, the speculation is obviously based

upon using the applicants' own claimed invention as a template - rather than a

legitimate analysis of what is actually taught to those of only ordinary skill in the art by

the cited prior art documents." (Remarks, Page 14, ¶ 1) The Examiner disagrees.

Although the Examiner still contends that the aspects of the claimed invention are

obvious in consideration of the rejection of 10/30/2009, a further explanation will be

supplied. It is fact that Vysotsky teaches speaker dependent speech recognition along

with grammars as was noted in the rejection. It is also fact that Vysotsky performs voice

verification (indication of the identity of the user). To further show that the claimed

aspects are obvious, the Examiner refers to Vysotsky, column 8, lines 50-65, ...*During

speaker dependent speech recognition, speaker dependent garbage models are used

for out of vocabulary rejection as well as word spotting. These garbage models are built

on-line and, in the exemplary embodiment, modified every time the user's directory of

names is changed...The grammar is shared by both the speech recognition and training

process used to generate the templates used for speech recognition. The grammar can

be modified by adjusting path probabilities to achieve various levels of word spotting

and rejection...* The grammar is modified each time the user's directory is changed to

accommodate for variances in pronunciation of the names in the directory. The

grammar is specifically modified such that the input speech from the recognizer may
better recognize the word from the grammar instead of another word in the grammar by
adjusting the path probabilities to achieve various levels of word spotting and rejection.
Although Vysotsky explicitly teaches this, it is further noted that Kunninen teaches a
directory as well (Page 18, lines 22-29, database) where it further would have been
obvious to someone of ordinary skill in the art to update select a user-dependent
speech grammar for the speech recognizer such that a person's name in the database
will be correctly recognized based on the user's particular pronunciation. The argument
is not persuasive.

10.   Applicant further argues "The Examiner's assertion that it would have been
obvious to combine these teachings "to provide a way for a person who receives a
message to know who sent the message to verify the identify of the caller" relies solely
upon Vysotsky at 5:45-50, but as previously noted, this passage deals with the arbiter
254 being coupled to a call completion and feature activation circuit 256 by a line 257
and by a voice verification circuit 255. While this does apparently provide an
arrangement for voice verification to be performed selectively (when, for security
purposes, it is important to verify the identity of a caller before responding to a particular
command), that does not offer any teaching or suggestion of the subject matter of
claims 24 and 34. In particular, the claims require that an indication of the identity of the
user who generated the audio stream be received - followed by selection of a user-
dependent speech grammar for use by a speech recognition process in dependence
upon the received user identity. There is nothing in Vysotsky that teaches using a

different speech grammar in a speech recognition process in dependence upon a

received user identity. If the Examiner continues to believe differently, then it is

respectfully requested that this teaching be more particularly identified." (Remarks,

Pages 15-16) The Examiner disagrees. See response in section 7.

Specification

11.　　The title of the invention is not descriptive. A new title is required that is clearly

indicative of the invention to which the claims are directed.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of
matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the
conditions and requirements of this title.

12.　　Claim 27 of the claimed invention is directed to non-statutory subject matter. The

claim recites "A data storage medium…" where the data storage medium is not clearly

defined in the specification only pertain to a statutory embodiment of the invention (no

carrier waves). The Examiner suggests using "A non-transitory data storage medium…"

to overcome this rejection as is shown in the 1351 OG 212 dated Feb 23rd 2010.

Appropriate correction is required.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

13.   Claims 19-23, 25-28, 29-33, and 35-36 are rejected under 35 U.S.C. 103(a) as

being unpatentable over Kinnunen et al. (WO 03/100372) in view of Brems. (US Patent

6505161)

As per claim 19, Kinnunen teaches the method comprising:

using a push-to-talk communication device during generation of an audio stream;

(Page 3, lines 15-26, …*When using the PoC feature, the user pushes the transmission*

key in the earpiece of his terminal equipment…)

in response to a user pressing a button on the push-to-talk communication

device and starting to talk, receiving at a router server in a communications network an

audio stream containing an utterance which includes an indication of an intended

receiver of the audio stream;   (Fig. 5 (31.1-31.2), Page 17, lines 25-34, …*it is*

possible for the user A', Bf, Cf to choose such individual users from his group, to whom

he addresses the transmission just by uttering, for example, the keyword stored as the

identifier corresponding to the user intended to be the recipient. In this way the user

may transmit private messages directly only to this certain user of his choice…)

buffering the received audio stream;   (Page 17, lines 1-10, …*Buffering*

of packets and timing/sequencing of transmissions to recipients is controlled with the

PoC server 31.1, 31.2…)

determining, if possible, an intended receiver of the audio stream in dependence upon the recognized utterance; and   (Page 17, lines 25-34, …*it is possible for the user A', B', C' to choose such individual users from his group, to whom he addresses the transmission just by uttering, for example, the keyword stored as the identifier corresponding to the user intended to be the recipient. In this way the user may transmit private messages directly only to this certain user of his choice…*)

if an intended receiver was determined, transmitting, to the determined intended receiver, the audio stream containing the utterance including the indication of the intended receiver of the audio stream, using a half-duplex communications service provided by a packet-switched network, such that no separate dialing phase is required. (Page 17, lines 25-34, …*it is possible for the user A', B', C' to choose such individual users from his group, to whom he addresses the transmission just by uttering, for example, the keyword stored as the identifier corresponding to the user intended to be the recipient. In this way the user may transmit private messages directly only to this certain user of his choice…* page 18, lines 15-20 describe the packet switched network. Further see, Page 17, lines 10-34 and Page 18, lines 22-29 where no separate dialing phase is needed.)

Kinnunen fails to specifically teach, but Brems teaches:

performing in a speech recognition server in said communications network a speech recognition process on the received audio stream to recognize the utterance

contained therein;   (Kinnunen, Page 13, lines 18-25, teaches a local speech

recognition process but fails to specifically address performing speech recognition at a

server (a distributed system). Brems, Fig. 2 (56, 64) and column 9, lines 3-7 teaches a

speech recognition server for distributed recognition.)

     It would have been obvious to someone of ordinary skill in the art at the time of

the invention to combine Brems with Kinnunen to provide distributed speech recognition

where different speech models may be applied for different input (local) device to

account for variances in background/environment noise between devices. (Brems,

column 2, lines 9-35)

As per claim 20, claim 19 is incorporated and Kinnunen teaches:

     indicating the one or more possible intended receivers to a user; and receiving a

selection signal from the user indicating the one or more determined possible intended

receivers to which said audio stream should be transmitted.   (Page 17 line 34 to

Page 18 line 2, *...The feature of the described kind can of course also be activated by

hand as a menu selection, but in certain conditions it is more natural to do this by

talking...*)

As per claim 21, claim 20 is incorporated and Kinnunen teaches:

     wherein the indicating step further comprises generating an audio speech prompt

corresponding to the one or more possible intended receivers; and outputting the

generated audio speech prompt to the user.   (Page 18, lines 22-29, *...If the*

VRE module finds the receiving party in its database, a confirmation of an established

form is given, which indicates a successful choice of voice. The confirmation may be, for

example, a short beep sound or a repetition of the keyword to the user…)

As per claim 22, claim 19 is incorporated and Kinnunen teaches:

wherein when the determining step determines a plurality of intended receivers, the audio stream is transmitted to each of the determined receivers using a group call function of the half- duplex communications service.   (Page 18, lines 6-13, … *The word "group", for example, may be stored as a keyword referring to the whole group…)*

As per claim 23, claim 19 is incorporated and Kinnunen teaches:

wherein the speech recognition process is performed only on a portion of the received audio stream when the intended recipient is indicated at the beginning of the audio stream.   (Page 17, lines 11-24, Kinnunen teaches that a password can be provided prior to the sentence. Although the password is not necessarily the intended user, it would have been obvious that the name or identifier of the intended recipient would be provided at the beginning of the audio stream because a person of common sense addresses a target individual before speaking to them. Also, the speech recognition (VRE) is only performed on a portion of the audio stream because the received audio is continuous (noise/speech), but the speech recognition is only performed on speech (a portion).)

As per claim 25, claim 19 is incorporated and Kinnunen teaches:

further comprising the steps of receiving a speech recognition activation signal from a user, wherein the speech recognition and determining steps are performed in dependence on the receipt of such a signal.   (Fig. 4a, there is a voice activity detection to check for voice regions, if there is no voice present the recognition engine isn't used.)

As per claim 26, claim 19 is incorporated and Kinnunen teaches:

monitoring audio streams transported by the half-duplex communications service; (Page 14, lines 21-23, ...*Hereby the audio signal is processed with the VAD or VRE function 23 (416) during the transmission*...)

performing a speech recognition process on the monitored audio streams to determine the respective utterances contained therein; and   (Page 14, lines 21-23, ...*Hereby the audio signal is processed with the VAD or VRE function 23 (416) during the transmission*...)

if it is determined that a predetermined utterance is contained in any of the audio streams, signaling that the half-duplex communications service should cease transporting the audio stream.   (Page 15, lines 15-11,...When the VRE function 23 finds the word over in the talk signal, the conclusion can be drawn that the intention is to end the transmission...)

Claims 27-28 are rejected for the same reasons as claim 19. The additional limitation of

a data storage medium is inherent in Kinnunen, Page 7, lines 19-22, *...When*

implemented entirely on a software basis without any 20 additional equipment or

components installed in the terminal equipment, the VOX feature as a combination of

VAD and VRE functions significantly reduces variable costs... Software requires a

computer readable medium to function.

Claims 29-33, and 35-36 are rejected for the same reasons as claims 19-23, 25-26. Fig.

1 teaches a system to function as the terminal equipment for the method as described

in Kinnunen.

14.   Claims 24 and 34 are rejected under 35 U.S.C. 103(a) as being unpatentable

over Kinnunen et al. (WO 03/100372) in view of Brems. (US Patent 6505161) and

further in view of Vysotsky et al. (US Patent # 5832063).

As per claim 24, claim 19 is incorporated and Kinnunen and Brems fails to specifically

teach, but Vysotsky teaches:

        receiving an indication of the identity of a user who generated the message;

(Vysotsky, column 5, lines 45-50, *...the arbiter 254, in turn, is coupled to a call*

completion and feature activation circuit 256 by a line 257 and by a voice verification

circuit 255. Using this arrangement, voice verification is performed selectively when, for

security purposes, it is important to verify the identity of a caller before responding to a

particular command..., A voice identification of the user teaches a means for receiving

an indication of the identity of the user in the instant application.)

grammar selection means for selecting a user-dependent speech grammar for

use by the speech recognition process in dependence on the identity of the user.

(Vysotsky, column 8, lines 31-35, *...The speaker dependent speech recognition*

process, like the speaker independent speech recognition process, is based on hidden

Markov models (HM) with the use of grammars..., The speaker dependent model is

based on grammars where a voice verification ability has been disclosed in Vysotsky,

column 5, lines 45-50. Thus, there would be a grammar selection means for selecting a

user-dependent speech grammar dependent for the specific user if voice verification

was performed on the individual.)

It would have been obvious to someone of ordinary skill in the art at the time of

the invention to combine Vysotsky with Kinnunen and Brems to provide a way for a

person who receives a message to know who sent the message to verify the identity of

a caller. (Vysotsky, column 5, lines 45-50)

Claim 34 is rejected for the same reasons as claim 24. Fig. 1 teaches a system to

function as the terminal equipment for the method as described in Kinnunen.

Conclusion

15.   The prior art made of record and not relied upon is considered pertinent to

applicant's disclosure. Refer to PTO-892, Notice of References Cited for a listing of

analogous art.

16.   Applicant's amendment necessitated the new ground(s) of rejection presented in

this Office action.   Accordingly, **THIS ACTION IS MADE FINAL**.   See MPEP

§ 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37

CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the date of this final action.

17.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to GREG A. BORSETTI whose telephone number is

(571)270-3885, (FAX: 571-270-4885). The examiner can normally be reached on

Monday - Thursday (8am - 5pm Eastern Time).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, RICHEMOND DORVIL can be reached on 571-272-7602. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Greg A. Borsetti/
Examiner, Art Unit 2626

/Talivaldis Ivars Smits/
Primary Examiner, Art Unit 2626   5/3/2010